Published in final edited form as:
Biometrics. 2012 December ; 68(4): 1157–1167. doi:10.1111/j.1541-0420.2012.01774.x.

Spatial-Temporal Modeling of the Association between Air Pollution Exposure and Preterm Birth: Identifying Critical Windows of Exposure

Joshua Warren[1], **Montserrat Fuentes**[2], **Amy Herring**[1], and **Peter Langlois**[3]

Joshua Warren: joshuawa@email.unc.edu

[1]Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7420, U.S.A.

[2]Department of Statistics, North Carolina State University, Raleigh, North Carolina 27695-8203, U.S.A.

[3]Texas Department of State Health Services, Austin, Texas 78714-9347, U.S.A.

Summary

Exposure to high levels of air pollution during the pregnancy is associated with increased probability of preterm birth (PTB), a major cause of infant morbidity and mortality. New statistical methodology is required to specifically determine when a particular pollutant impacts the PTB outcome, to determine the role of different pollutants, and to characterize the spatial variability in these results. We develop a new Bayesian spatial model for PTB which identifies susceptible windows throughout the pregnancy jointly for multiple pollutants ($PM_{2.5}$, ozone) while allowing these windows to vary continuously across space and time. We geo-code vital record birth data from Texas (2002–2004) and link them with standard pollution monitoring data and a newly introduced EPA product of calibrated air pollution model output. We apply the fully spatial model to a region of 13 counties in eastern Texas consisting of highly urban as well as rural areas. Our results indicate significant signal in the first two trimesters of pregnancy with differe0nt pollutants leading to different critical windows. Introducing the spatial aspect uncovers critical windows previously unidentified when space is ignored. A proper inference procedure is introduced to correctly analyze these windows.

Keywords

Air pollution; Bayesian analysis; MCMC; Preterm birth; Probit regression; Spatial model

1. Introduction

The association between pollution exposures and harmful pregnancy outcomes is becoming more evident due to the increased number of recent studies investigating the relationship. In 2005, Šrám et al. concluded, after an extensive literature review, that evidence linking air pollution with adverse birth outcomes exists. The authors found sufficient evidence that a causal relationship can be inferred between ambient air pollution exposure and low birth

Correspondence to: Joshua Warren, joshuawa@email.unc.edu.

Supplementary Materials

Web Appendices, Tables, and Figures referenced in Sections 1, 3, and 4 are available with this paper at the Biometrics website on Wiley Online Library.

weight but suggested that more information was necessary to examine the effect of different pollutants and to determine the most vulnerable periods of the pregnancy.

The relationship between preterm birth (PTB), a delivery occurring before 37 completed weeks of gestation, and air pollution exposure is not as well understood. Šrám et al. concluded that there was not yet enough information to draw any general conclusions regarding this relationship but did suggest that more studies were justified given the supporting information. More recently, Ritz et al. (2007) conducted a case control study of women in Los Angeles and modeled the probability of PTB while linking it with air pollution. They consistently found an estimated increase in vulnerability to PTB for higher exposures to $PM_{2.5}$ and carbon monoxide in the first trimester. In 2006, Leem et al. carried out a similar study on PTB using data from Incheon, Republic of Korea, and also observed evidence identifying the first trimester as the most vulnerable period of the pregnancy for the considered pollutants.

These results are important in firmly establishing the link between PTB and air pollution in general. We extend these results by more specifically identifying the critical time periods during the pregnancy when exposure to air pollution is particularly harmful in the context of PTB. The exposure to harmful pollutants during the pregnancy is typically handled through trimester or monthly averages and fit separately using multiple models, including separate models for different pollutants. Conducting the analysis in this way is inefficient and does not allow the joint identification of specific periods across the entire pregnancy in a continuous manner. It is also common for these standard statistical models to ignore the spatial aspect of the data. A single region is analyzed with individual pollution exposures determined by closest monitor matching and the risk effects assumed to be constant over the domain.

In this paper we introduce a single Bayesian probit regression model with spatially and temporally varying effects for PTB. The model simultaneously handles multiple pollutants and jointly models time periods that account for the entire span of pregnancy for each woman in the study while allowing these effects to vary over the geographic domain. Ambient exposures to pollutants are more accurately estimated through spatial-temporal modeling and pregnancy specific prediction of climatic and pollution variables in initial stages.

Fitting the model in the Bayesian setting allows for a more flexible solution to obtaining parameter estimates and associated uncertainty measures in this situation. The typical frequentist analysis requires the maximization of the likelihood function, which is difficult to even specify given the complexity of the dependence structure present in the data. Using a hierarchical model and Markov chain Monte Carlo (MCMC) techniques simplifies this process and allows us to carry out the usual inference in a more efficient way.

This model allows us to identify the specific critical windows of exposure over the entire pregnancy that lead to a higher probability of PTB and to see how these windows potentially change across the spatial domain. It also gives more insight into how different pollutants affect the pregnancy in different ways. We fit this model using a dataset of hospital births in Texas that has not been analyzed before in relation to air pollution, the standard pollution monitoring data, and a recently introduced form of pollution estimate data provided by the Environmental Protection Agency (EPA).

The birth dataset covers 2002–2004 and includes parental demographic and birth outcome information on all births in Texas during these years. We successfully geo-code residence at delivery for a majority of the women in the dataset as well. The geo-coding process is explained in Section 2.1. Weekly averages of ambient $PM_{2.5}$ and ozone concentrations,

based on each woman's location and dates of pregnancy, are estimated using modeled and predicted Air Quality System (AQS) monitoring data (Air Quality System). Next, the weekly exposures are estimated using the EPA provided pollution data known as the Statistically Fused Air and Deposition Surfaces data (FSD) (McMillan et al., 2010). These data represent daily pollution surface estimates, with standard errors, on a grid over the eastern or conterminous US, depending on spatial resolution. To our knowledge, this is the first time the FSD data are used in an environmental health project. The results from fitting the model in Harris County using the different pollution datasets are compared and proper inference is carried out for the indicated critical windows. The fully spatial model, which allows the risk effects to be correlated spatially, is then applied to a heterogeneous region of 13 counties in Texas using the predicted AQS data. This region provides a number of rural and urban counties where the critical windows are estimated and compared.

Our newly introduced model allows us to simultaneously handle the exposure from multiple pollutants in a continuous manner throughout the entire pregnancy while accounting for spatial changes in the risk effects. The model is able to better identify susceptible windows of importance for PTB temporally and spatially. This work helps to increase the evidence linking air pollution and birth outcomes while extending the results for preterm births.

In Section 2, we describe the data used in the analysis while Section 3 introduces the statistical model. In Section 4, the statistical model is applied to the Texas birth dataset and the results are presented along with model diagnostics and prior sensitivity analysis. We close in Section 5 with the conclusions and further discussions. Derivations are presented in Web Appendix A and sample code is shown in Web Appendix B.

2. Data Description and Preparation

2.1 Health Data

The dataset we analyze consists of full birth records for all births in Texas from 2002–2004. To be included in the analysis, each infant must have been delivered in 2002–2004 as a singleton, live birth. The mother must have resided in Texas at the time of delivery and had no previous live births. Additionally, at least some demographic information must have been available on the infant and mother. All of the included data come from vital records (birth certificates). Available pregnancy information includes date of birth, sex, birth weight, date of last menstrual period, clinical estimate of gestational age, and parental information such as age, birthplace, race and ethnicity, and education level. In the analysis, preterm delivery is a binary variable defined as a delivery occurring before 37 completed weeks of gestation, based on the clinical estimate of gestational age.

We geo-coded the data in order to include location information for the pregnancies. The geo-coding process was carried out by the Geographic Information System group at the Texas Department of State Health Services (TDSHS). The process attaches coordinate information to the address files in the vital records data using street and address location software. Cleaned addresses were linked to latitudinal and longitudinal coordinates through an automated process. An interactive matching process was used for addresses that could not be linked in this way. Addresses were linked to a central street segment if the streets were entirely contained in a single US census block group. Records were linked to the nearest intersection when the street number was not given. The ZIP code centroid of a residence was used to link records that could not be linked using any of the previous methods.

We assume that from the date of conception to the date of birth a woman in the analysis remains at her residence at delivery. A study in Texas from 1995–2002 found that 68% of pregnant women did not move between date of conception and date of birth (Nuckols et al.,

2004). Of the women who did move, 49% moved to a location with the same water supply source. Another study in New York State (excluding NYC) found that 16.5% of women changed addresses but moved on average only 10.4 miles (median: 2.6) (Chen et al., 2010). While this suggests that a majority of pregnant women remain near their residence at delivery through the entire pregnancy, there is still potential to misclassify the assigned exposures for the women who did move. This is likely due to the spatial variability of the $PM_{2.5}$ pollutant as moving a short distance can potentially impact the amount and composition experienced.

2.2 Pollution Data

The AQS monitoring data are available in Texas from 2002–2004. The AQS is a collection of ambient air pollution data from thousands of monitoring stations throughout the US. The data are collected by the EPA, state, and local air pollution agencies and are used by the Office of Air Quality Planning and Standards and others in a number of air quality management functions (AQS: Basic information).

The maximum daily 8-hour average ozone values (parts per million (ppm)) are used in the analysis. These values were typically collected daily by monitors in Texas from 2002–2004. To attain the National Ambient Air Quality Standards (NAAQS) standard for ozone, the annual fourth-highest daily maximum 8-hour concentrations, averaged over 3 years, measured at each monitor within an area must not exceed 0.075 ppm (NAAQS).

The daily average $PM_{2.5}$ values (micrograms per cubic meter (ug/m^3)) are obtained as well. These values were typically collected every three to six days by Texas monitors. To attain the NAAQS standard for $PM_{2.5}$, the 98th percentile of 24-hour concentrations, averaged over 3 years, at each monitor within an area must not exceed 35 g/m3 (NAAQS).

The FSD data are a new EPA product representing $PM_{2.5}$ (daily average) and ozone (daily 8-hour maximum) pollution surface estimates on a grid over the entire Community Multiscale Air Quality (CMAQ) 12km × 12km and 36km × 36km spatial resolutions for 2002–2004. These data represent an alternative form of pollution information to the standard AQS monitoring data based on combining multiple data sources into a single pollution estimate. The process used to create the FSD product calibrates the CMAQ numerical model output using monitoring data from the AQS. The CMAQ estimates are potentially biased with respect to the monitoring data and the FSD data attempt to correct for this bias using a modification of techniques developed by Fuentes and Raftery (2005).

The process used to create the FSD data relies on sound statistical models and techniques to combine multiple sources of information into a single estimate. Some disadvantages associated with the FSD data include a lack of information regarding the shape of the posterior distribution of the true pollution process (only mean and standard deviation estimates given) as well as a potential bias due to not taking into account the change of support problem caused by having data on different spatial resolutions (Fuentes, 2009).

The CMAQ modeling system combines information from a number of scientific areas in order to model multiple air quality issues simultaneously, including tropospheric ozone, fine particles, toxics, acid deposition, and visibility degradation. CMAQ provides gridded estimates of ozone, particulates, toxics, and acid deposition at different resolutions across the US using this expertise in air quality modeling and atmospheric science (CMAQ overview).

In the analysis the exposure to the considered pollutants represents the ambient concentrations at the residence at delivery. We do not have information regarding daily

activities that is needed to account for more specific exposures such as those experienced at work, during travel, and inside the home. Currently indoor exposure is not regulated by the government.

2.3 Climatic Data

Nation-wide daily meteorological data are available from the National Climate Data Center. We obtain the daily average temperature, dewpoint, cloud cover, windspeed, and sea level pressure from monitors in Texas, 2002–2004.

3. The Statistical Model

We introduce a hierarchical framework for analyzing the association between the preterm delivery outcome and air pollutant concentrations. In the first stage we introduce a model for the climatic data. In Stage 2 we adapt the statistical model for the pollution data, originally introduced by Fuentes and Raftery (2005), based on the AQS monitoring observations and modeled climatic data. In the third stage we develop the probit regression model for PTB.

In the full model, the stages are treated somewhat separately with the posterior predictive distribution (ppd) from one stage entering the next stage as a prior distribution. This directional Bayesian technique is known as a "cut of feedback" approach which includes both practical and computational benefits (Lunn et al., 2009). In a fully joint hierarchical Bayesian modeling framework, health data could provide information about the ppd of the weather variables, which is often viewed as an undesirable intuitive property. Therefore, estimates are obtained separately at each stage and the corresponding uncertainty is captured at the final stage of the hierarchical model. Web Figures 1 and 2 display graphical outlines of the modeling process.

In Stage 1 of the analysis a model for the climatic data is introduced and samples from the ppd are be obtained using the Bayesian kriging technique at locations/dates of interest. The general model for the weather observations has the form

$$\mathbf{W}(\boldsymbol{s},t)=\mu_{\mathrm{w}}(\boldsymbol{s},t)+\mathbf{e}_{\mathrm{w}}(\boldsymbol{s},t)+\boldsymbol{\varepsilon}_{\mathrm{w}}(\boldsymbol{s},t), \quad (1)$$

where $\mathbf{W}(\boldsymbol{s}, t)$ represents the vector of weather values at location $\boldsymbol{s}$ and time t; $\boldsymbol{\mu}\mathrm{w}(\boldsymbol{s}, t)$ represents the large scale spatial and temporal trend of the data; $\mathbf{e}\mathrm{w}(\boldsymbol{s}, t)$ is modeled as a spatially and temporally correlated zero mean Gaussian process; and $\boldsymbol{\varepsilon}\mathrm{w}(\boldsymbol{s}, t)$ represents an independent white noise process, serving as a nugget effect for the process. The excellent space-time coverage (>130 active monitors) and overall smoothness of the climatic data allows us to separate Stages 1 and 2 and ease the computational burden without losing information. The pollution data were not needed to provide information about the climatic variables in this case.

In Stage 2 we introduce the model for the pollution data. We work with the $PM_{2.5}$ pollutant as an example. We assume that the true underlying $PM_{2.5}$ process, $Z_1(\boldsymbol{s}, t)$, is unobservable without measurement error, and we model it using the weather variables from the initial stage such that

$$Z_1(\boldsymbol{s},t)=\mu_1(\boldsymbol{s},t)+\mathbf{W}^{(p)}(\boldsymbol{s},t)^T\delta_1+\mathrm{e}_1(\boldsymbol{s},t). \quad (2)$$

The $\mathrm{e}_1(\boldsymbol{s}, t)$ errors are assumed to be from a zero mean Gaussian process, spatially and temporally correlated and the $\boldsymbol{\delta}_1$ vector represents coefficients relating the weather process to the pollution. The ppd of the weather variables from Stage 1 becomes the prior distribution in Stage 2 for the predicted weather process, $\mathbf{W}^{(p)}(\boldsymbol{s}, t)$. The large scale space-

time deterministic trend is represented by $\mu_1(\boldsymbol{s}, t)$. We assume that observations from the AQS monitoring system, $\tilde{Z}_1(\boldsymbol{s}, t)$, represent unbiased estimates of the true underlying $PM_{2.5}$ process at location $\boldsymbol{s}$ and time t, with some associated measurement error,

$$\tilde{Z}_1(\boldsymbol{s}, t) = Z_1(\boldsymbol{s}, t) + \varepsilon_1(\boldsymbol{s}, t). \quad (3)$$

The $\varepsilon_1(\boldsymbol{s}, t)$ errors are assumed to be independent and normally distributed with mean zero and some constant variance representing the measurement error associated with the pollution monitors. Values of $Z_1(\boldsymbol{s}, t)$ are simulated from the ppd, $f(\mathbf{Z}_1|\tilde{\mathbf{Z}}_1)$, at each woman's location during the relevant pregnancy window. These ppd samples determine the prior distribution of the pollution exposure used by the Stage 3 PTB health model.

Using probit regression we introduce a continuous pollution exposure model in Stage 3 for the Texas birth data. The choice of the probit link results in conjugacy in the model which is given as

$$\begin{gathered} Y_i|\beta, \theta \overset{ind}{\sim} \text{Bernoulli}\{p_i(\beta, \theta)\}, \\ p_i(\beta, \theta) = \text{the probability that birth } i \text{ results in preterm birth, and} \\ \Phi^{-1}\{p_i(\beta, \theta)\} = \mathbf{x}_i^T\beta + \sum_{j=1}^{2} \sum_{w=1}^{min(ga_i, 36)} \theta\{j, w, \boldsymbol{B}(\boldsymbol{s}_i)\} Z_j^{(p)}\{\boldsymbol{s}_i, t_i(w)\} \end{gathered} \quad (4)$$

where $\Phi^{-1}(.)$ represents the inverse cumulative distribution function of the standard normal distribution and 'ga_i' is the gestational age (weeks) for birth i. The $\boldsymbol{B}(\boldsymbol{s}_i)$ term represents the block or more general region of interest containing location $\boldsymbol{s}_i$. For our application we allow $\boldsymbol{B}(\boldsymbol{s}_i)$ to represent the county within TDSHS health service region six, shown in Web Figure 3, containing location $\boldsymbol{s}_i$. This choice is explained further in Section 4.5. In some situations $\boldsymbol{B}(\boldsymbol{s}_i) = \boldsymbol{s}_i$ which becomes computationally expensive but is theoretically accounted for under this generalized model formulation.

The $\mathbf{x}_i$ vector contains parental covariates and other confounders of interest. This includes an intercept term, parental age group, race, education, and seasonality information. We consider six age groups for the mothers: 10–19, 20–24, 25–29, 30–34, 35–39, and 40+. We include two age groups for the fathers: < 50 and 50+. For both mothers and fathers we consider White (non-Hispanic), Black (non-Hispanic), Hispanic, and Other as the four race/ethnic groups in the analysis. To control for education level, the number of years of completed education is fit using a cubic B-spline with three degrees of freedom. To account for seasonality we include the average temperature, predicted in Stage 1, from the day of birth using a cubic B-spline with four degrees of freedom.

The $\theta\{j, w, \boldsymbol{B}(\boldsymbol{s})\}$ parameters are pollutant specific, spatially and temporally-varying coefficients. They represent the effect of the concentration of air pollutant j at pregnancy week w (corresponding to calendar week $t_i(w)$) at location $\boldsymbol{s}$ within region $\boldsymbol{B}(\boldsymbol{s})$ on the probability of PTB for woman i. We consider two pollutants, ozone and $PM_{2.5}$, in the analysis. The predicted pollution exposure process for pollutant j on calendar week $t_i(w)$ at location $\boldsymbol{s}_i$ is represented by $Z_j^{(p)}\{\boldsymbol{s}_i, t_i(w)\}$ where the prior is determined by the ppd obtained in Stage 2. The summation $\sum_{j=1}^{2} \sum_{w=1}^{min(ga_i, 36)} \theta\{j, w, \boldsymbol{B}(\boldsymbol{s}_i)\} Z_j^{(p)}\{\boldsymbol{s}_i, t_i(w)\}$ contains a different number of terms for each woman due to the fact that women have different gestational ages. This prevents pollution exposure after the birth of the child from affecting the probability that a birth results in a preterm outcome.

The $\theta\{j, w, \boldsymbol{B}(\boldsymbol{s})\}$ parameters are modeled using a Gaussian process prior distribution with mean zero and a specific covariance structure based on the belief that $\theta\{j, w, \boldsymbol{B}(\boldsymbol{s})\}$ parameters closer in space, time, and associated with the same pollutant are more highly correlated. The full prior process of the $\boldsymbol{\theta}$ vector is $\boldsymbol{\theta} \sim \text{MVN}(\mathbf{0}, \boldsymbol{\phi}_0\boldsymbol{\Sigma})$, where entries of $\boldsymbol{\phi}_0\boldsymbol{\Sigma}$ are given by $\text{Cov}[\theta\{j, w, \boldsymbol{B}(\boldsymbol{s})\}, \theta\{j', w', \boldsymbol{B}(\boldsymbol{s}')\}] =$

$$\phi_0 \exp\{-\phi_1 \|\boldsymbol{B}(\boldsymbol{s}) - \boldsymbol{B}(\boldsymbol{s}')\| - \phi_2 |w - w'| - \phi_3 I(j \neq j')\}. \quad (5)$$

The covariance parameters, ϕ_0; ϕ_1; ϕ_2; ϕ_3, are each > 0 and $I(j \neq j')$ takes value one for different pollutants ($j \neq j'$) and is zero if $j = j'$. Including this exponential covariance structure provides a relatively simple parameterization that still allows separate degrees of shrinkage across air pollutants (j), pregnancy weeks (w), and locations ($\boldsymbol{B}(\boldsymbol{s})$). It also allows us to overcome the multicollinearity that we introduce by considering an effect for each week for multiple pollutants.

Appropriate prior distributions are chosen for the parameters to complete the model. The ϕ_1, ϕ_2, and ϕ_3 parameters are given different combinations of vague uniform and gamma priors and the results are compared. Relatively uninformative inverse gamma and uniform priors are chosen for the overall variance parameter, ϕ_0, and results are compared. The $\boldsymbol{\beta}$ parameters are given independent, vague yet proper normal prior distributions.

3.1 Inference for Critical Windows

Using the available MCMC output from the fitting of the health model in Stage 3 allows a straight forward method for conducting inference on the critical pregnancy windows. The model has the ability to individually identify particular week effects, $\theta\{j, w, \boldsymbol{B}(\boldsymbol{s})\}$, which are larger than zero by providing associated posterior mean estimates and credible intervals for each effect. After visually inspecting plots of these effects across pregnancy weeks (along with credible intervals), the question becomes whether or not a particular span of these individually positive effects jointly represent a critical window of exposure during the pregnancy. By combining the resulting model fit output we can estimate $P\left[\theta^{\{j,\boldsymbol{B}(\boldsymbol{s})\}}_{(window)} > \mathbf{0} | \mathbf{y}_{obs}\right]$, where $\theta^{\{j,\boldsymbol{B}(\boldsymbol{s})\}}_{(window)} = [\theta\{j, w_1, \boldsymbol{B}(\boldsymbol{s})\}, \ldots, \theta\{j, w_n, \boldsymbol{B}(\boldsymbol{s})\}]^T$ and $w_1, \ldots, w_n$ ($n \leq 36$) are the weeks included in the specific window of interest at location $\boldsymbol{B}(\boldsymbol{s})$ and for pollutant j. The specific weekly span of the window is selected by analyzing the individual effects at a certain location and for a selected pollutant. In this way we can jointly estimate the probability that the critical windows, which are suggested by the individual positive effects, are also jointly impactful. Results from this analysis are found in Section 4.1.

4. Application to Texas Birth Data

We begin in Stage 1 by separately modeling the climatic variables in the Bayesian setting after applying appropriate normalizing transformations. The deterministic trend in (1) includes a combination of latitude/longitude and seasonal information which is found to provide an adequate fit of the large scale structure.

The structure for the space-time error in (1) consists of independent, identical spatial processes across time such that $e\mathrm{w}(\boldsymbol{s}, t) = w_t(\boldsymbol{s})$ where $\mathbf{w}_t = \{w_t(\boldsymbol{s}_1), \ldots, w_t(\boldsymbol{s}_n)\}^T \overset{iid}{\sim} \text{MVN}(\mathbf{0}, \Sigma_w)$. The particular choice of the spatial covariance structure, Σ_w, varies for each climatic variable and is chosen based on standard model selection techniques. Typical choices are the exponential and spherical isotropic models. Assuming an isotropic correlation structure leads to

$$\text{Cov}\{e_w(\boldsymbol{s},t), e_w(\boldsymbol{s}',t')\} = \text{Cov}\{w_t(\boldsymbol{s}), w_{t'}(\boldsymbol{s}')\} = \begin{cases} \sigma_w^2 \rho(\|\boldsymbol{s}-\boldsymbol{s}'\|) & t=t' \\ 0 & t \neq t' \end{cases}$$ where ρ (.) represents the particular isotropic correlation function and ‖.‖ is the Euclidean distance between the spatial locations. This error structure choice consistently outperformed other space-time structures and represents a common choice in the literature (Peng and Bell, 2010; Banerjee, Carlin, and Gelfand, 2004). Other tested structures include separate, independent effects for date and location, $e_w(\boldsymbol{s}, t) = \alpha(t) + w(\boldsymbol{s})$, and spatially independent time series at each location, $e_w(\boldsymbol{s}, t) = \alpha_s(t)$. Web Table 8 details the specific spatial covariance function used for each weather variable along with posterior summaries of the associated covariance parameters.

In Stage 2 we separately model the $PM_{2.5}$ and ozone AQS responses in the Bayesian setting after applying normalizing transformations to each variable. The deterministic trend in (2) for each variable includes latitude/longitude, seasonal, and location/date specific predicted weather covariate information. The climatic variables are given multivariate normal prior distributions where the mean vectors and covariance matrices are determined by the ppd samples from Stage 1. This multivariate normal approximation is determined to be adequate based on the multivariate normality of the modeled climatic variables and an investigation of the ppd samples from Stage 1. Web Figure 4 shows examples of ppd histograms and normal quantile-quantile plots for two of the climatic variables.

The most appropriate structure for the space-time error structure in (2) is once again shown to be $e_j(\boldsymbol{s}, t) = w_t(\boldsymbol{s})$. For ozone, the corresponding Σ_W is chosen to follow the isotropic exponential structure $(\rho(\|\boldsymbol{s}-\boldsymbol{s}'\|) = \exp(-\phi\|\boldsymbol{s}-\boldsymbol{s}'\|))$. The anisotropic exponential covariance structure $(\rho(\boldsymbol{s},\boldsymbol{s}') = \exp\left(-\sum_{k=1}^{2}\phi_k|s_k - s_k'|^{p_k}\right))$ is chosen for the $PM_{2.5}$ variable where $\boldsymbol{s} = (s_1, s_2)$. Web Tables 9 and 10 display posterior summaries for all parameters in the ozone and $PM_{2.5}$ models respectively.

We implement a computationally efficient algorithm which allows us to handle the large spatial-temporal domains of interest in the analysis. We obtain samples from the ppd, $f(\boldsymbol{Z}_j|\tilde{\boldsymbol{Z}}_j)$, for every day in 2002–2004 on the 36km × 36km CMAQ grid. The number of necessary predictions are cut from N × 3 × 365 to G × 3 × 365 by working on the grid, where N is the total sample size of the birth dataset (≈ 450,000) and G is the total number of grid locations in Texas (541). Computationally this method is very attractive but it also introduces an assumption about the pollution process that requires further investigation. It assumes that $Z_j(\boldsymbol{s}, t) = Z_j\{\boldsymbol{A}(\boldsymbol{s}), t\} + \varepsilon(\boldsymbol{s}, t)$ where $Z_j\{\boldsymbol{A}(\boldsymbol{s}), t\}$ is the pollution value at grid cell location $\boldsymbol{A}(\boldsymbol{s})$ with location $\boldsymbol{s}$ lying within $\boldsymbol{A}(\boldsymbol{s})$. The $\varepsilon(\boldsymbol{s}, t)$ error is the random deviation arising from a white noise process with variance σ_A^2 representing the spatial variability occurring within a grid cell. For a reasonably small grid cell area σ_A^2 should be negligible. To verify this assumption empirically, random grid cells are chosen and $\tilde{Z}_j(\boldsymbol{s}, t)$ is predicted at various dates and locations within the cell and compared to the respective predictions of $\tilde{Z}_j\{\boldsymbol{A}(\boldsymbol{s}), t\}$. Web Figure 6 displays the plots from this analysis for $PM_{2.5}$ and ozone from a grid cell in Harris County. These plots and simple linear regression results suggest that σ_A^2 is indeed negligible and this computationally efficient algorithm is appropriate. Therefore, ppd samples from $f(Z_j\{\boldsymbol{A}(\boldsymbol{s}), t\}|\tilde{\boldsymbol{Z}}_j)$ are summarized as in Stage 1 and used to create prior distributions for the Stage 3 health model. Ppd histograms and normal quantile-quantile plots are shown for each pollutant in Web Figure 5.

4.1 Harris County Results: AQS Model Fit

Table 1 summarizes the various models considered in the analysis. Model 1A results are presented first. The spatial aspect is ignored due to the smaller geographic size of the county by allowing $\boldsymbol{B}(\boldsymbol{s})$ in (4) to represent the county containing location $\boldsymbol{s}$. As of July 1, 2009, Harris County had the third largest county population in the United States with 4,070,989 estimated people (American Fact Finder). It includes the city of Houston, which provides a large amount of heterogeneity to our study population.

We begin by examining the included covariate results. We then analyze the most influential weeks identified by the model on the probability of PTB for each pollutant. The final dataset we analyze includes 32,170 observations with results based on 10,000 draws from the posterior distribution after a burn in period of 10,000 draws.

Table 2 shows the posterior summaries for the covariates included in the analysis. The estimated effect represents the increase (or decrease) in z-score for a one-unit increase in the explanatory variable. An increased z-score leads directly to an increased probability of PTB. Therefore, a positive effect implies an increase in the probability of PTB with a similar interpretation for a negative effect.

The range and average values of the Monte Carlo (MC) error for the means are also given. The MC error is an estimate of the standard error of the mean and is calculated using the batched means method detailed by Roberts (1996).

The maternal covariates appear to have the greatest effect on PTB. When compared with White (Non-Hispanic) mothers, Black (Non-Hispanic) mothers have a higher probability of PTB in general. White (Non-Hispanic) fathers are more likely to father preterm children when compared to fathers who claim Other as their race/ethnic group. Mothers age 35 and older are at a higher risk of having a PTB outcome with 20–24 appearing to be the ideal maternal age in terms of PTB. Males are more likely than females to be born prematurely. Overall, these results agree with previously established epidemiological results.

The major advantage of the newly introduced model lies in its ability to identify the critical windows of exposure during the pregnancy. Figure 1 shows this feature graphically. The posterior median of each weekly pollution effect is plotted against pregnancy week along with the respective 95% credible intervals for each pollutant. From these plots it is clear which periods during the pregnancy lead to an increase in the probability of PTB.

The susceptible window for higher preterm probabilities covers the middle of the first trimester through the middle of the second trimester for the $PM_{2.5}$ pollutant. Week 14 has the largest estimated effect with a posterior mean of 0.0327 (MC error: 0.0002, SD: 0.007). This means a one unit increase in the standardized pollution exposure for week 14 leads to an increase in z-score of 0.0327 on average. Increased $PM_{2.5}$ exposure in weeks 4–22 increases the risk of PTB as each of the respective credible intervals do not include zero.

The ozone results differ, with high exposure in the very early weeks of pregnancy appearing to increase the probability of PTB. Week 1 has the most drastic effect in terms of ozone exposure with a posterior mean of 0.0207 (MC error: 0.0003, SD: 0.008). A similar interpretation exists for this effect. Increased ozone exposure in weeks 1–5 increases the risk of PTB based on the associated credible intervals.

The graphical results in Figure 1 suggest possible critical windows of exposure for each pollutant which need to be analyzed jointly in order to conduct proper inference. For $PM_{2.5}$ the relevant window appears to stretch from week 4 to 22. It is obvious that individually the

weekly effects are likely larger than zero since their respective credible intervals do not include zero and by combining the output we estimate $P\left\{\theta^{(1)}_{(window)}>\mathbf{0}|\mathbf{y}_{obs}\right\}$. This describes the relevant weekly effects jointly and allows inference for the critical windows to be carried out as opposed to individual weekly analyses alone. This quantity is estimated to be 0.945 (MC error: 0.004) for the $PM_{2.5}$ results with $\theta^{(1)}_{(window)}=\theta^{(1)}_{(4-22)}$. For ozone $P\left\{\theta^{(2)}_{(1-5)}>\mathbf{0}|\mathbf{y}_{obs}\right\}$ is estimated to be 0.971 (MC error: 0.002). Therefore the critical windows which are evident graphically are also impactful jointly.

The results suggest that early in the pregnancy, the developing fetus may be most vulnerable. Exposure to air pollution early in the first trimester may interfere with the delivery of oxygen and nutrients to the fetus. This exposure may affect the placental development during the early stages of pregnancy as well. There is also evidence suggesting that the exposure to air pollution may trigger inflammation, leading to PTB. The exact explanation for how the pollution affects the fetus has not yet been identified. Some studies show that ultra fine particles can enter the mother's lungs and penetrate the lung barriers, entering the bloodstream. These particles can then travel to the organs, such as the brain and placenta, and may cause problems for the fetus (Ritz and Wilhelm, 2008).

4.2 Model Diagnostics

The immediate benefits of our introduced model are seen when the results are directly compared to results from more standard statistical methods. In Model 2 we consider a common birth outcome model with first and second trimester averages from multiple pollutants included jointly, failing to account for the introduced multicollinearity (Bobak, 2000). A more naive approach is applied in Model 3 by including all of the weekly exposures from multiple pollutants in a large multiple probit regression model, also ignoring the present correlation between weeks and pollutants. This provides a crude way of carrying out the analysis as we would expect the parameter estimates to be similar with their respective standard errors possibly inflated due to the multicollinearity. We fit both of these models using probit regression in the Bayesian setting for comparison purposes. The graphical results from Models 1A, 2, and 3 are shown in Figure 2. Complete posterior summaries of Model 2 and Model 3 covariate parameters are displayed in Web Tables 11 and 12 respectively.

While the plots from the standard methods do indicate some weeks/trimesters whose credible intervals do not include zero, we miss the majority of the values that are clearly seen when using our model. Model 3 is unable to uncover the true state of the windows of exposure. The uncertainty associated with the risk estimates is much larger when compared to Model 1A as the plots from both models are shown on the same scale. Model 2 does not include a fine enough time scale to be as informative as our newly introduced Model 1A.

The deviance information criterion (DIC) is used to carry out a more formal comparison of the models (Spiegelhalter et al., 2002). The DIC is based on the posterior distribution of the deviance statistic, $D(\boldsymbol{\gamma}) = -2 \log f(\mathbf{y}|\boldsymbol{\gamma}) + 2 \log h(\mathbf{y})$, where $f(\mathbf{y}|\boldsymbol{\gamma})$ represents the likelihood of the observed data given the $\boldsymbol{\gamma}$ vector of parameters and $h(\mathbf{y})$ is some standardizing function of the data alone. The posterior expectation of the deviance, $\bar{D}$, is used to describe the fit of the model to the data while the effective number of parameters, p_D, is used to describe the complexity of the model. DIC is then defined as $DIC = \bar{D} + p_D$, with lower values of DIC representing a better model fit. The DIC criterion clearly favors Model 1A (DIC: 17246.3, p_D: 42.4) when compared with Model 2 (DIC: 17282.8, p_D: 28.3) and Model 3 (DIC: 17289.1, p_D: 96.2) as differences of more than seven in DIC rule out the model with the higher value.

Using techniques described by Dey and Chen (2000) we investigate the overall adequacy of the fitted model by performing posterior predictive comparisons. The following diagnostics ensure that our model fits the data relatively well.

We first define the observation-level Pearson residual discrepancy measure as

$D_i(y_i;\beta,\theta)=\frac{\{y_i - p_i(\beta,\theta)\}^2}{p_i(\beta,\theta)\{1 - p_i(\beta,\theta)\}}$ where $p_i(\boldsymbol{\beta}, \boldsymbol{\theta})$ is given in (4). We consider the total Pearson residual discrepancy measure, $D(\mathbf{y};\beta,\theta)=\sum_{i=1}^{N} D_i(y_i,;\beta,\theta)$ since the overall performance of the model is of interest. Values of the discrepancy measure are simulated from the ppd, $f\{D(\mathbf{y}_{new}; \boldsymbol{\beta}, \boldsymbol{\theta}) |\mathbf{y}_{obs}\}$, and also from the observed data distribution, $f\{D(\mathbf{y}_{obs}; \boldsymbol{\beta}; \boldsymbol{\theta}) |\mathbf{y}_{obs}\}$, where $\mathbf{y}_{obs}$ represents the vector of observed outcomes and $\mathbf{y}_{new}$ represents the vector of simulated outcomes from the ppd, $f(\mathbf{y}_{new}|\mathbf{y}_{obs})$. The samples from these respective distributions are then compared to assess the overall fit of the model to the data. Two different Bayesian exploratory data analysis methods are used to compare these samples.

First, we compare the distributions of interest graphically. Web Figure 7 shows the boxplots of samples from $f\{D(\mathbf{y}_{obs}; \boldsymbol{\beta}, \boldsymbol{\theta}) |\mathbf{y}_{obs}\}$ and $f\{D(\mathbf{y}_{new}; \boldsymbol{\beta}, \boldsymbol{\theta}) |\mathbf{y}_{obs}\}$ respectively. When the model provides an adequate fit of the data, we expect these two distributions to be similar and therefore the plots to look very much alike. It is evident from the plots that Models 1A and 2 appear to provide an adequate fit of the data while the boxplots from Model 3 appear to differ much more, indicating poor fit.

Next, we estimate $P\{D(\mathbf{y}_{new}; \boldsymbol{\beta}; \boldsymbol{\theta}) \geqslant D(\mathbf{y}_{obs}; \boldsymbol{\beta}, \boldsymbol{\theta}) |\mathbf{y}_{obs}\}$. This quantity is often referred to as the Bayesian p-value and is useful in determining model adequacy (Meng, 1994). The estimates for Models 1A and 2 are 0.130 and 0.172 respectively (MC errors: 0.005 and 0.007) while for Model 3 the estimate is 0.016 (MC error: 0.002). This provides more evidence to suggest that Models 1A and 2 provide an adequate fit while Model 3 again performs poorly.

4.3 Sensitivity to Priors

The overall results for the pollution and covariate effects do not change when we use different prior distributions for the covariance hyper-parameters. In Web Tables 1–4, the posterior summaries for the covariance parameters and DIC estimates are shown for Models 1A–1D. These different combinations of priors lead to very similar estimates of the covariance parameters which lead to similar windows of exposure being identified as critical in terms of PTB probability.

4.4 Harris County Results: FSD Model Fit

In Model 4 we use the FSD data to construct the pollution prior distributions. Use of the FSD data replaces Stages 1 and 2 of the presented statistical model. The dataset of births being analyzed remains the same as for the presented Model 1A fit results with only the prior distributions for the pollution differing. The graphical results from Models 1A and 4 are shown in Figure 1. Complete posterior summaries for Model 4 parameters are displayed in Web Tables 5 and 13.

We compare the results using a sensitivity analysis approach. Validating the FSD product is difficult as we only obtain the finished product of posterior means and standard deviations estimates. The posterior distributions for $\boldsymbol{\beta}$ are very similar for both of the model fits as each of the individual credible intervals agree in terms of impact. The ϕ_0 parameter estimates and credible intervals are also nearly identical for both fits. The temporal smoothness parameter, ϕ_2, is much smaller for Model 4, indicating a smoother process in time. The ϕ_3 parameter is

also estimated to be smaller for Model 4. This implies a greater amount of information sharing across the windows from the different pollutants. These changes in the estimates of the covariance parameters for Model 4 result in changes in the risk effect estimates. The smaller ϕ_3 parameter estimate leads to similar risk effect estimates for both of the pollutants, also decreasing the overall variability of the estimates. None of the risk effects now have credible intervals which fail to include zero for any of the weeks. The FSD data for the $PM_{2.5}$ pollutant are less reliable in this situation given that they rely heavily on CMAQ because of the lack of daily monitoring data. Air quality numerical model data handle spatial variability very well but are known to have difficulty accounting for the temporal variability (Hogrefe et al., 2001). Model 1A results are therefore more reliable and lead to critical windows which go undetected once the FSD product is used in Model 4.

4.5 Fully Spatial Results: TDSHS Health Service Region Six

In order to implement the fully spatial version of the Stage 3 health model we extend the spatial domain to include the 13 counties from TDSHS health service region six, shown in Web Figure 3. We assign $\boldsymbol{B}(\boldsymbol{s}_i)$ to represent the location of the county containing location $\boldsymbol{s}_i$. The specific location of each county is determined by the center of gravity of the births within each county. TDSHS health service region six provides a good mix of urban and rural counties which are useful for investigating possible changes in risk effects due to different compositions of the pollutants across these counties. Web Figure 3 also displays the residence at delivery of the women in this region included in the analysis and shows the rural/urban areas more clearly. We choose to work at the county level after examining within Harris County and finding that the risk effects do not vary over the smaller spatial domain. These results are shown in Web Figure 8. Including 13 counties in the analysis brings the sample size to 43,607 and the displayed results are based on 5,000 draws from the posterior distribution after a burn in period of 5,000 draws.

We fit two models to the 13 county dataset in order to investigate the benefits of considering the spatial correlation between the risk effects. Model 5 ignores the spatial variability that exists between the risk effects by adapting (5) so that

$$\text{Cov}\left[\theta\{j, w, \boldsymbol{B}(\boldsymbol{s})\}, \theta\{j', w', \boldsymbol{B}(\boldsymbol{s}')\}\right] = \phi_0 \exp\{-\phi_2|w - w'| - \phi_3 I(j \neq j')\}. \quad (6)$$

This structure assumes spatial independence of the risk effects across the 13 counties. Next, we fit Model 6 which accounts for the present spatial correlation by using the full covariance structure in (5). Complete posterior summaries for Model 5 are shown in Web Tables 6 and 14 with Model 6 results shown in Web Tables 7 and 15. The DIC criterion clearly favors Model 6 (DIC: 23857.4, p_D: 61.4) over Model 5 (DIC: 23920.4, p_D: 80.6). Web Figure 9 shows the output for Brazoria, Waller, and Montgomery counties. The decrease in the variability of the posterior distributions of the risk effects as a result of accounting for the spatial correlation is easily seen by comparing the plots (shown on same scale). The graphical results from Model 6 suggest that the risk effects are consistent over the region. This spatial smoothing agrees with the graphical output from Model 5 which indicates similar trends of the risk effects in each county. Web Figure 9 also shows how impactful weeks are able to be identified once the fully spatial model is applied and are missed when spatial independence is assumed. The level of temporal smoothing for the non-spatial plots is increased due to the lack of an overall signal when space is ignored.

Figure 3 shows the estimated posterior means and standard deviations of the average combined risk effects plotted across the 13 counties for the first trimester of pregnancy from Models 5 and 6. The first trimester average combined risk effect at location $\boldsymbol{s}$ has the form

$\bar{\theta}_1\{\boldsymbol{B}(\boldsymbol{s})\}=\frac{1}{13}\sum_{w=1}^{13}[\theta\{1,w,\boldsymbol{B}(\boldsymbol{s})\}+\theta\{2,w,\boldsymbol{B}(\boldsymbol{s})\}]$. This quantity gives some insight into the combined impact from both pollutants averaged over the first trimester on the probability of PTB at various spatial locations. The posterior standard deviation plots show even more clearly the reduction in variation that occurs when we account for the spatial correlation. The posterior distributions of the risk effects for Galveston, Liberty, and Walker counties have 95% credible intervals which include zero for the Model 5 results while being larger than zero once Model 6 is applied. Model 6 is able to identify these positive effects which are missed when space is ignored in Model 5.

The posterior means plot from Model 6 gives better insight into the true spatial relationship which exists between the risk effects and suggests that the eastern section of the region has higher estimated effect sizes than the western section. A possible explanation for the observed spatial pattern of the posterior means is the age breakdown within each county. The four counties in the west which have the smallest estimated effects each have at least 13.4% of their population greater than or equal to 65 years in age with Colorado county leading the way with 18.6%. The overall rate for the entire state is 10.2% while none of the other nine counties have a rate higher than 11.1% (American Fact Finder). Along with the smaller sample sizes already present in these more rural counties, this age breakdown may be leading to the observed lack of signal in the west.

5. Discussion/Conclusion

Using our model, we are able to simultaneously characterize the effect of exposure to multiple pollutants on the PTB outcome in a continuous manner, throughout the entire span of pregnancy. We are also able to account for the existing spatial correlation between the risk effects and investigate their possible changes across space. A proper inference method is introduced to correctly investigate the identified critical windows of exposure. The benefits of the model are obvious when compared to the standard PTB models found in the literature as well as simplified multiple regression techniques. These models fail to identify specific windows during the pregnancy and are clearly outperformed by our model. Implementing the fully spatial model decreases the variability in the risk effect estimates while giving a clearer picture of the spatial trend of the effects. These results further build the evidence supporting the link between air pollution and PTB while extending our knowledge regarding the specific periods during the pregnancy that have the greatest impact in terms of PTB.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank Tom Luben, US EPA, for his input which contributed greatly to this work. We thank the National Science Foundation (Fuentes DMS-0706731, DMS-0934595), the EPA (Fuentes, R833863), and the National Institutes of Health (Fuentes, 5R01ES014843-02) for partial support of this work.

References

Air Quality System. USEPA; 2011 Aug 12. Web. Feb. 2012.

American Fact Finder. US Census Bureau; n.d. Web. Feb. 2012.

AQS: Basic information. Air Quality System. USEPA; 2010 Nov 5. Web. Feb. 2012.

Banerjee, S.; Carlin, BP.; Gelfand, AE. Hierarchical modeling and analysis for spatial data. Boca Raton, FL: Chapman & Hall/CRC; 2004.

Bobak M. Outdoor air pollution, low birth weight, and prematurity. Environmental Health Perspectives. 2000; 108:173–176. [PubMed: 10656859]

Chen L, Bell EM, Caton AR, Druschel CM, Lin S. Residential mobility during pregnancy and the potential for ambient air pollution exposure misclassification. Environmental Research. 2010; 110:162–168. [PubMed: 19963212]

CMAQ overview. Community Multiscale Air Quality Model. CMAS Center and the Center for Environmental Modeling for Policy Development at UNC-CH; n.d. Web. Feb. 2012.

Dey, DK.; Chen, MH. Bayesian model diagnostics for correlated binary data. In: Dey, D.; Ghosh, S.; Mallick, B., editors. Generalized linear models: a Bayesian perspective. New York, NY: Marcel Dekker, Inc.; 2000. p. 320-327.

Fuentes M. Statistical issues in health impact assessment at the state and local levels. Air Quality, Atmosphere & Health. 2009; 2:47–55.

Fuentes M, Raftery AE. Model evaluation and spatial interpolation by Bayesian combination of observations with outputs from numerical models. Biometrics. 2005; 61:36–45. [PubMed: 15737076]

Hogrefe C, Rao ST, Kasibhatla P, Hao W, Sistla G, Mathur R, et al. Evaluating the performance of regional-scale photochemical modeling systems: part II-ozone predictions. Atmospheric Environment. 2001; 35:4175–4188.

Leem JH, Kaplan BM, Shim YK, Pohl HR, Gotway CA, Bullard SM, et al. Exposures to air pollutants during pregnancy and preterm delivery. Environmental Health Perspectives. 2006; 114:905–910. [PubMed: 16759993]

Lunn D, Best N, Spiegelhalter D, Graham G, Neuenschwander B. Combining MCMC with sequential PKPD modelling. Journal of Pharmacokinetics and Pharmacodynamics. 2009; 36:19–38. [PubMed: 19132515]

McMillan NJ, Holland DM, Morara M, Feng JY. Combining numerical model output and particulate data using Bayesian space-time modeling. Environmetrics. 2010; 21:48–65.

Meng X. Posterior predictive p-values. The Annals of Statistics. 1994; 22:1142–1160.

National Ambient Air Quality Standards (NAAQS). Air and Radiation. USEPA; 2011 Nov 8. Web. Feb. 2012.

Nuckols, JR.; Langlois, P.; Lynberg, ML.; Luben, T. Linking geographic water utility data with study participant residences from the National Birth Defects Prevention Study. Denver, CO: American Water Works Association Research Foundation; 2004.

Peng RD, Bell ML. Spatial misalignment in time series studies of air pollution and health data. Biostatistics. 2010; 11:720–740. [PubMed: 20392805]

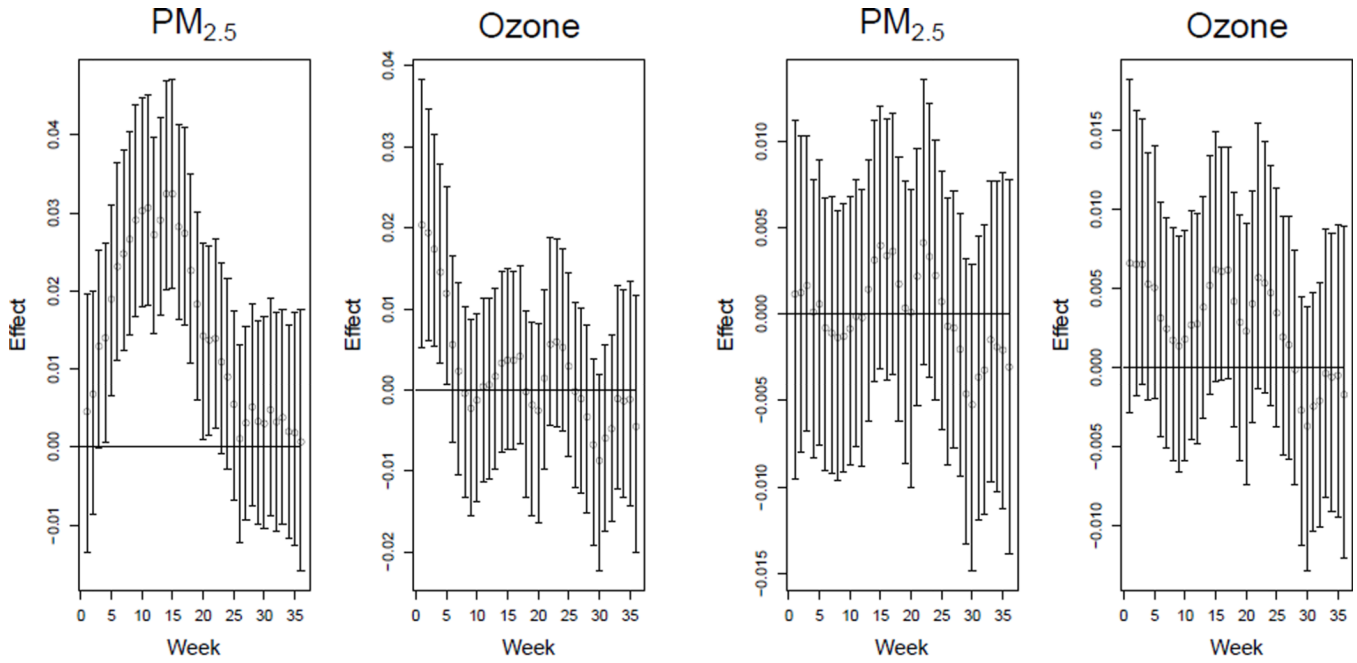
Ritz B, Wilhelm M. Air pollution impacts on infants and children. Southern California Environmental Report Card. 2008 Web. May. 2011.

Ritz B, Wilhelm M, Hoggatt KJ, Ghosh JKC. Ambient air pollution and preterm birth in the environment and pregnancy outcomes study at the University of California, Los Angeles. American Journal of Epidemiology. 2007; 166:1045–1052. [PubMed: 17675655]

Roberts, GO. Markov chain concepts related to sampling algorithms. In: Gilks, WR.; Richardson, S.; Spiegelhalter, DJE., editors. Markov chain Monte Carlo in practice. London: Chapman and Hall; 1996.

Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society Series B-Statistical Methodology. 2002; 64:583–616.

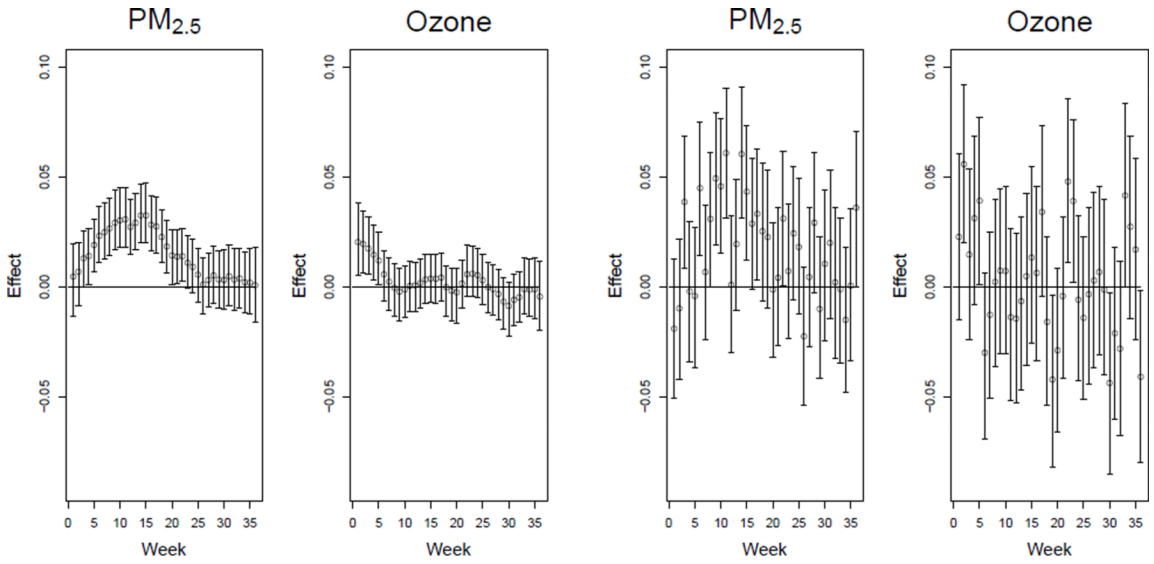
Šrám RJ, Binkova BB, Dejmek J, Bobak M. Ambient air pollution and pregnancy outcomes: a review of the literature. Environmental Health Perspectives. 2005; 113:375–382. [PubMed: 15811825]



(a) Model 1A: AQS Results.

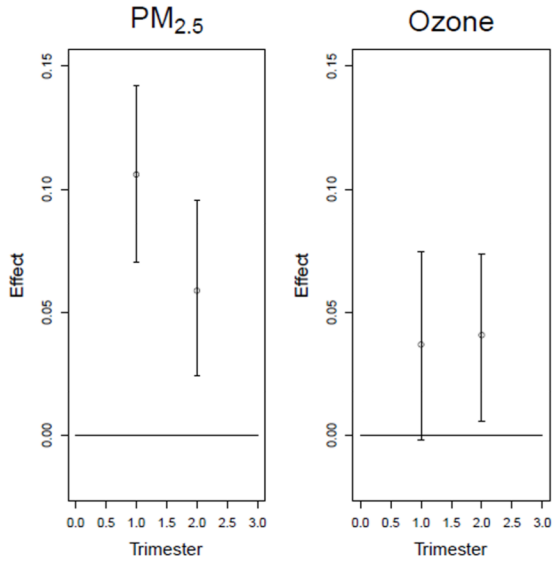
(b) Model 4: FSD Results.

Figure 1.
Susceptible windows of exposure results from Model 1A (left) and Model 4 (right). Posterior medians and 95% credible intervals are displayed.



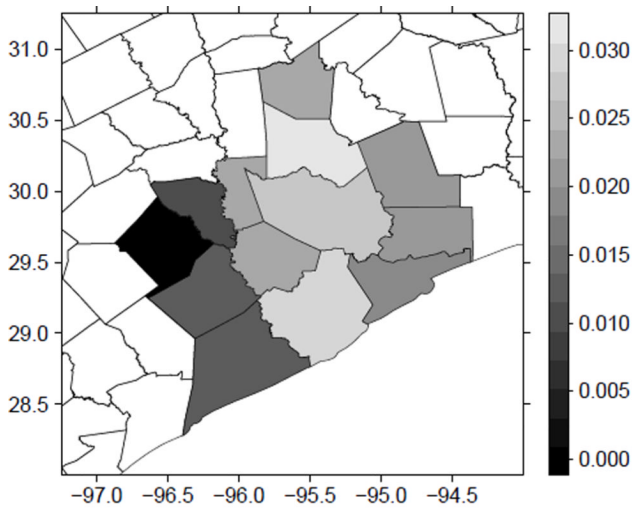
(a) Model 1A: Newly Introduced Model.

(b) Model 3: Multiple Probit Model.

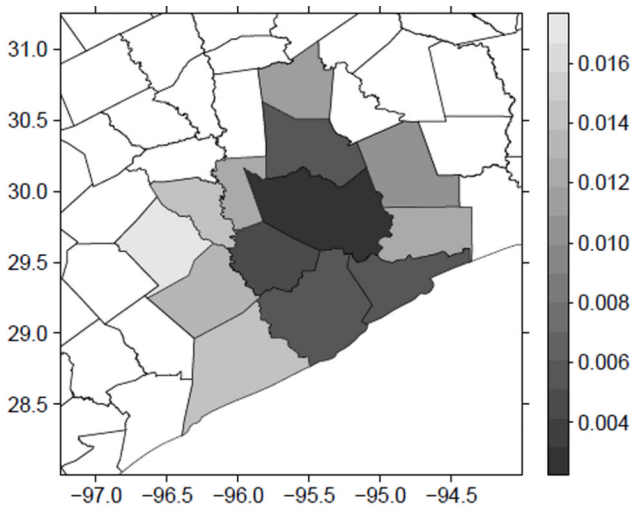


(c) Model 2: Trimester Average Model.

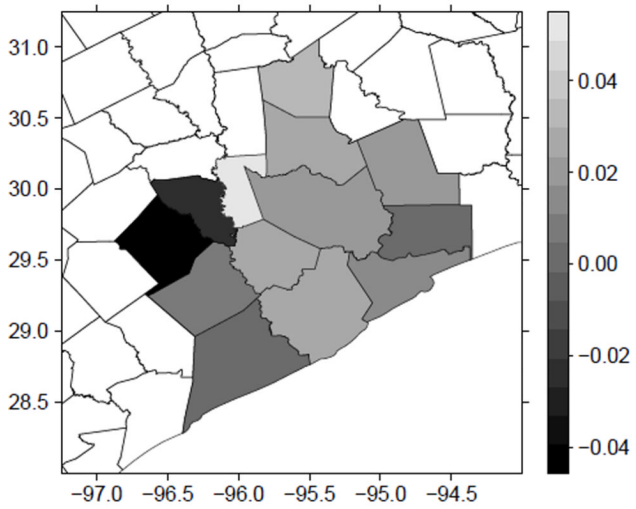
Figure 2.
Susceptible windows of exposure results from Model 1A, Model 2, and Model 3. Posterior medians and 95% credible intervals are displayed.



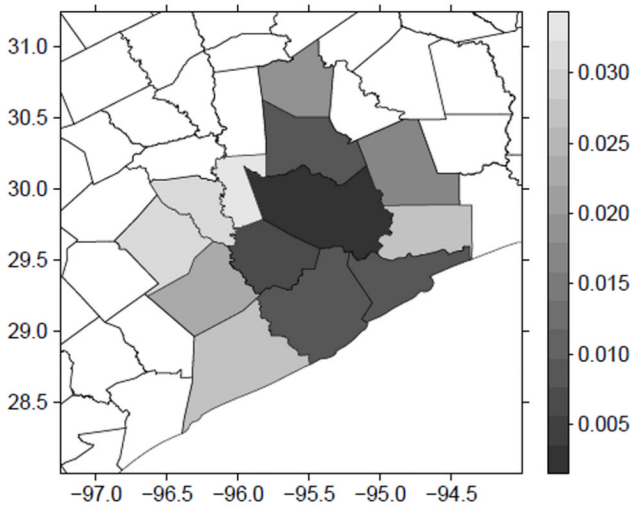
(a) Posterior Means: Spatial (Probit Scale).



(b) Posterior Standard Deviations: Spatial.



(c) Posterior Means: Non-Spatial (Probit Scale).



(d) Posterior Standard Deviations: Non-Spatial.

Figure 3.
Posterior means and standard deviations of first trimester averages of the combined $PM_{2.5}$ and ozone risk effects by county for Model 6 (top row) and Model 5 (bottom row) model fits.

Table 1

Bayesian probit regression models considered in the analysis. All models control for the same covariate information.

Model ID	Description
	Harris County: 2002–2004
1A–1D **	Non-spatial version of the PTB health model utilizing the full weather and pollution stages in the modeling process.
2	Standard epidemiological PTB model jointly considering trimester 1 and 2 pollution exposure averages for $PM_{2.5}$ and ozone (estimated using the full weather and pollution stages) as fixed model inputs.
3	Multiple probit regression model jointly considering all weekly $PM_{2.5}$ and ozone exposure averages (estimated using the full weather and pollution stages) throughout the pregnancy as fixed model inputs.
4	Non-spatial version of the PTB health model utilizing the FSD pollution product (replaces the weather and pollution stages).
	TDSHS Region 6: 2002–2004
5	Spatially independent version of the PTB health model where the risk parameters are allowed to vary spatially but the spatial correlation is ignored (full weather and pollution stages are utilized).
6	Fully spatial version of the PTB health model where the spatial correlation structure between the risk parameters is implemented (full weather and pollution stages are utilized).

(**) Models 1A–1D differ only by the prior distributions assigned to the covariance hyper-parameters. Web Tables 1–4 detail these prior distribution settings and posterior summaries.

Table 2

Included covariate results for Model 1A.

			Percentiles		
Covariate	**Mean**	**SD**	**0.025**	**0.50**	**0.975**
Intercept **	−1.515	0.279	−2.065	−1.513	−0.979
Maternal Race					
Black vs. White **	0.149	0.067	0.018	0.149	0.280
Hispanic vs. White	−0.043	0.038	−0.117	−0.044	0.033
Other vs. White	0.086	0.070	−0.053	0.086	0.227
Paternal Race					
Black vs. White	0.009	0.065	−0.118	0.010	0.135
Hispanic vs. White	−0.008	0.039	−0.084	−0.009	0.070
Other vs. White **	−0.190	0.073	−0.337	−0.189	−0.046
Maternal Age Group					
20 – 24 vs. 10 – 19 **	−0.068	0.033	−0.132	−0.069	−0.003
25 – 29 vs. 10 – 19	−0.021	0.036	−0.091	−0.021	0.052
30 – 34 vs. 10 – 19	0.044	0.041	−0.036	0.045	0.126
35 – 39 vs. 10 – 19 **	0.127	0.052	0.026	0.127	0.229
⩾ 40 vs. 10 – 19 **	0.308	0.090	0.128	0.309	0.478
Paternal Age ⩾ 50 vs. < 50	−0.044	0.128	−0.302	−0.041	0.204
Maternal Education					
Basis Spline 1	−0.396	0.348	−1.053	−0.401	0.303
Basis Spline 2	0.185	0.172	−0.140	0.180	0.529
Basis Spline 3	0.022	0.204	−0.360	0.018	0.434
Paternal Education					
Basis Spline 1	0.160	0.307	−0.429	0.153	0.786
Basis Spline 2	−0.003	0.150	−0.290	−0.005	0.296

			Percentiles		
Covariate	**Mean**	**SD**	**0.025**	**0.50**	**0.975**
Basis Spline 3	−0.055	0.177	−0.389	−0.059	0.305
Female vs. Male Baby **	−0.069	0.021	−0.110	−0.069	−0.027

The (**) items have 95% credible intervals which do not include zero. The MC error for the means ranged from 0.0005 to 0.0077 with an average value of 0.0027.